

Rethinking the assumptions of chi-squared and fisher's exact tests

Almost always when we are confronted with a contingency table, we summon to mind the chi-squared (χ^2) test, which was developed more than a century ago.¹ Like most classic statistical tests, the χ^2 test can be conducted with pencil and paper, so long as the raw data are summarized; it is seemingly straightforward and even considered one of the easiest analyses, although this is not true. Several complex modern statistical methods may analyze contingency tables in a more sophisticated way, but they do not add much practical benefit in most situations, especially in the healthcare field. Thus, the χ^2 test has been one of our closest friends for a long time, as it deserves to be.

Despite the χ^2 test's handiness, we should make sure that the χ^2 test is an appropriate method for the data on hand before running it by checking its assumptions, which is what we actually do. If the assumptions do not hold, we search for an alternative method, praying there is one. Luckily, we have Fisher's exact test, especially for 2x2 tables, as the immediate next option in our armory.²

Then, what exactly are the assumptions of the χ^2 test? Commonly, the following two are reckoned to be important and widely known. First, each observation is independent; in other words, 'one observation for one subject'.³ Second, "no more than 20% of the expected counts are less than 5 and all individual expected counts are 1 or greater".⁴ The first assumption is relatively clear with any data and no calculation is required. If that does not hold, we can easily switch to McNemar's test. However, the second assumption cannot be checked with just a glance; it requires calculation. Look into the assumption again. The most important word here is *expected* counts. The raw numbers in the contingency table are *observed* counts au fond. The χ^2 test basically compares actual observed counts on the table with expected counts calculated with row totals and column totals. Therefore, only after calculating the expected counts of all cells can we decide whether the assumption is met or not. A very common mistake that leads us to set aside the χ^2 test arises from failing to distinguish expected and observed counts. More often than we imagine, researchers consider numbers in the contingency table as expected counts and, thus, regard encountering any number less than 5 in the table as a red light to stop the χ^2 test, which then forces researchers to switch into the fallback option, Fisher's exact test. As readers may well have understood by now, in the above scenario, dropping the χ^2 test is an unnecessary renunciation of our privilege to depend on a highly useful classic analysis; it is an unsanctioned under-utilization of a tool that should have been used.

At this moment, we should ask ourselves what the assumptions of Fisher's exact test are. Surprisingly, many researchers are not aware that there are assumptions for Fisher's exact test. If Fisher's exact is indeed an assumption-free test, then it would be a panacea for any contingency table-related tests; there is no need for agony regarding the assumptions of the χ^2 test. However, life is not that simple: In the realm of inferential statistics, all analysis methods require assumptions to be checked. Technically, to justify using Fisher's exact

Volume 6 Issue 1 - 2017

Heon-Jae Jeong,¹ Wui-Chiang Lee²¹The Care Quality Research Group, Chuncheon, Korea²Department of Medical Affairs and Planning, Taipei Veterans General Hospital & National Yang-Ming University School of Medicine, Taipei, Taiwan

Correspondence: Wui-Chiang Lee, Department of Medical Affairs and Planning, Taipei Veterans General Hospital & National Yang-Ming University School of Medicine, Taipei, Taiwan, Tel +886-2-28757120, Fax +886-2-28757200, Email leewuichiang@gmail.com

Received: June 12, 2017 | **Published:** June 22, 2017

test (especially the original Fisher's exact test), the investigator should know both row and column totals beforehand, which is rarely the case in the biomedical field.⁵ To illustrate, in a clinical trial for drug effectiveness, the researcher knows the number of patients who take and do not take the drug at the initiation of the study (beforehand), but she can never know the number of cured patients until the end of the trial.⁶ In contingency table terminology, she only knows total values in one margin but not the other at the beginning of the study. Strictly speaking, Fisher's exact test may not fit into this trial. In reality, meeting this requirement for Fisher's exact test may be more difficult than meeting that of the χ^2 test.⁷ Now that we know of and understand the assumptions of Fisher's exact test, albeit in somewhat theoretical and abstract terms, it would be fair to consider them at least as much as we do with the χ^2 test.

The above misunderstanding and ignorance have paved a vicious road of misuse, the road taken by too many people; that is, if a researcher misunderstands the assumptions of the χ^2 test—*especially the expected counts part*—she may inappropriately abandon the χ^2 test and apply Fisher's exact test. Because she does not know about the assumptions of Fisher's exact test, naturally she does not check them and just makes inferences and declares that the analysis is "fine." If you read academic articles carefully, you will be astonished at how easily you can find this kind of misuse expressed in the form of either underuse or overuse of these methodologies. By now, some might desperately ask what we can do if the assumptions of Fisher's exact test do not hold. Be relieved. Other alternatives can better suffice with the data. Only your volition to search for them is necessary, which can easily be done by searching the Internet or just opening your statistics textbook.

We all agree that there is no excuse for us not to check whether the assumptions of an analysis are met; indeed, we do check diligently. Thus, we want to emphasize the following: There is no excuse not to check whether we correctly understand the assumptions. It is

definitely an obligation and maybe a courtesy toward something that helps us every day. Ignorance or negligence may make us wrongfully accuse innocent analysis methodologies of misuse and break up our long-standing friendship with them⁸. Remember, the χ^2 and Fisher's exact tests do not make mistakes; we do.

Acknowledgement

None.

Conflict of interest

None.

References

1. Plackett RL. Karl Pearson and the chi-squared test. *International Statistical Review/Revue Internationale de Statistique*. 1983;51:59–72.
2. Gibbons JD. Fisher's exact test. *Wiley Stats Ref: Statistics Reference Online*. 1983.
3. Daniel WW, Wayne WD. *Biostatistics: A foundation for analysis in the health sciences*.
4. Yates D, Moore D, McCabe G. *The practice of statistics*. Freeman (Ed.) New York, USA. 1999.
5. Lantz K. Small-sample comparisons of exact levels for chi-squared goodness-of-fit statistics. *Journal of the American Statistical Association*. 1978;73(362):253–263.
6. Campbell I. Chi-squared and Fisher–Irwin tests of two-by-two tables with small sample recommendations. *Stat Med*. 2007;26(19):3661–3675.
7. Upton GJ. Fisher's exact test. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 1992;395–402.
8. Jeong H-J, Lee W-C. Ignorance or Negligence: Uncomfortable Truth Regarding isuse of Confirmatory Factor Analysis. *Journal of Biometrics & Biostatistics*. 2016;7(3):298.